Vision and fact: A critical essay on the growth literature

by

Professor Jan Fagerberg

Center for technology, innovation and culture, University of Oslo[i]

Abstract

This paper contains an overview and interpretation of the literature on economic growth. It is argued that, first and foremost, growth theory is about vision. It is concerned with questions such as: Where are we heading, and why? What are the options? Which actions are needed to arrive at the preferred state? Thus, growth theory is by its very nature deeply political.

Introduction

Many people think about growth theory as complicated mathematics. Others identify it with enormous amounts of numbers. And it is true; growth theory is about both of these things. And more. First and foremost, I will argue, growth theory is about vision. It is concerned with questions such as: Where are we heading, and why? What are the options? Which actions are needed to arrive at the preferred state? Thus, growth theory is by its very nature deeply political.

This became abundantly clear with the very advent of growth theory centuries ago, when writers such as Adam Smith and David Ricardo tried to convince people that stagnant growth and generally poor living conditions were not at all necessary, if only institutions and policies were geared towards allowing the capitalist machine to work at full speed. They argued forcefully that such changes in institutions and policies, although detrimental to the narrow interests of some stakeholders in the existing system, would be enormously beneficial to society as a whole. This "free market" optimism has been a central ingredient in many economists basic beliefs ever since.

Not all writers on growth were equally cheerful about the long run outcome of "free market" capitalism. For instance, writers such as Thomas Malthus and Karl Marx both held that an unregulated capitalist market system was deemed to stagnation, crises and – possibly - collapse, though for different reasons. Although mainstream (orthodox) economists generally rejected such ideas as faulty, economic developments in the first parts of this century actually seemed to confirm the crises-ridden character of capitalism. The view that capitalist growth if left to itself was not sustainable gained credibility as the crisis deepened and unemployment soared. The most central advocate of this view in the 1930s was of course John Maynard Keynes, emphasizing the importance of demand and the essential role of the government in managing the economy. His analysis – which was largely confined to the short run – was extended to the long run by among others Evsey Domar and Roy Harrod (so-called "post-Keynesian growth theory").[ii] These theories showed that long run growth with full employment was indeed possible but depended on extensive intervention by the government (especially with respect to income distribution). By the end of the Second World War the Keynesian view of the need for an active government had become widely shared among policy-makers and stakeholders in the Western world.

The 1950s and the return of orthodoxy

The Keynesian dominance in growth theory should not last long. There were several reasons for this. First, the conclusions of the Keynesian and post-Keynesian theories were in conflict with what economists had been taught to believe since the days of Adam Smith, i.e., that capitalism is a self regulating system that performs best when interference in markets is at a minimum. For many economists acceptance of this belief was (and in many quarters still is) the most important criterion for being an accepted member of the profession. Second, the times were changing. The 1930s had witnessed depression and unemployment. Now, in the 1950s, the economies of the West were running at full speed. Rather than economic misery the real threat to Western societies now was generally conceived to be the expansion of the state-led systems in Eastern Europe and Asia (the Soviet Union in particular). An economic theory that advocated extensive state intervention in the economy must have looked very odd from such a perspective.

Whatever the reason, several orthodox economists started to search for a new growth theory. The so-called neo-classical growth theory developed by Robert Solow and others in the 1950s proposed that long-run growth with full employment was indeed possible as long as market forces were allowed to operate freely. However, the strong assumptions behind this conclusion were severe. The theory rested on the idea of so-called "perfect competition", i.e., an economy with many competing firms, each of them too small to have a real impact on the market. In this idealized economy, economies of scale was ruled out by assumption, since this was believed to imply that large firms would have lower unit costs than smaller firms and, hence, would be able to drive the latter out of the market. Thus, constant returns to scale were imposed, i.e., that a 1 % increase in all inputs yields exactly 1 % growth in output. Technology was regarded as an exogenous force, a so called "public good", freely available to everyone free of charge. Thus, investments by firms in development of new technology were also ruled out, since such investment would carry no particular financial reward for the firm that undertook it.

In such an economy the only way for a firm or country to increase its productivity relative to its competitors would be to mechanize, i.e., increase the amount of capital per worker. But neoclassical economists also assumed that the rewards to such substitution of capital for labor gradually diminish as the amount of capital that each worker is set to operate increases. For a given level of technology this implies there exists a limit beyond which accumulation of capital per worker will not be profitable. When an economy has reached this limit, the amount of capital per worker has reached its "steady state" (equilibrium), and labor productivity will be constant (in the absence of exogenous technological change). Or, to put it differently, from this perspective the only source of long run productivity growth is exogenous technological advance.

As long as we accept these assumptions, this theory leads to an important prediction when applied to the global economy. Countries that differ in terms of initial productivity levels but not otherwise will converge towards the same level of productivity and the same rate of productivity growth. If countries differ also in other respects, such as the growth of population and the propensity to save, convergence towards the same growth of productivity will still be achieved, but long run productivity levels will differ. In the literature the latter is often called “conditional convergence”.

Hence, the neoclassical growth theory developed by Robert Solow and other in the 1950s implied a liberal and optimistic view on the prospects for global economic development. As long as market forces were allowed to operate freely, and other conditioning factors did not differ too much, everyone would all be equally well off in the long run.

Growth accounting

Although highly abstract, and based very strong the assumptions, at the time he neoclassical theory of growth was welcomed by empirical growth analysts who felt that it might give theoretical backing to their attempts to calculate the contribution of factor accumulation (labor and capital) to growth of GDP and productivity. This was so because, following to the theory, the prices of labor of capital (i.e., wages and profits) will reflect the contribution from of one additional unit of labor and capital, respectively, to GDP. Hence, in this case, one might actually use observed data on wages and profits to calculate the contribution from increased use of labor and capital to growth of GDP. This practice came to be known as “growth accounting”.[iii] In the 1950s, it was applied to historical data for the USA by Moses Abramovitz and others, and in the 1960s and later to selected OECD countries by Edward Denison. Over the years this methodology has also been applied to many individual countries.

In theory, in the absence of exogenous technological progress, the contributions from growth of labor and capital calculated in this way should add up to actual growth. If there is more growth than what can be accounted for by rising capital expenditure and more labor, the logic of the theory suggests that this should be seen as the contribution from exogenous technological progress. But it may of course also reflect other factors not taken into account by the theory. In practice, the first exercises carried out in this area showed that only a small part of actual growth could be attributed to growth of labor and capital, up to 80 % remained unexplained.

That the lion's share of actual growth had to be explained by exogenous technological progress and other unidentified sources was something many were not willing to accept right away. Various remedies were invented to improve on this rather embarrassing result. The first was to adjust the factors themselves by taking into account the changes in quality and composition. For instance, new vintages of capital, embodying the most recent technologies, were assumed to be more efficient than previous ones. To some extent this practice boiled down to no more than building the unexplained part of actual growth (exogenous technological progress etc.) into the factors themselves. As for labor, its quality was assumed to increase with increases in the educational level, the economic effect of which was normally calculated by observing the wage gaps between labor with different educational levels. Second, it was suggested to take into account additional factors such as economies of scale, investments in R&D, possible differences in productivity levels across countries and sectors and a host of other factors (crime, for instance). When these additional factors were taken into account the growth accountants were able to explain a much larger part of actual growth. There was only one problem, that the very existence of some of these additional factors actually contradicts the assumption of the theory on which the analysis was based. For instance, as shown previously, the theory explicitly assumes no economies of scale. Thus, if such additional factors are relevant, and the available evidence suggests that they are, what you need is a new theory!

In fact, Richard Nelson pointed the limitations of growth accounting out long ago.[iv] Growth accounting, he argued, is not a tested theory on growth, it is an analysis - or description - of a growth process based on certain assumptions, the validity of which have generally not been proven (or tested). Often – we may add - these assumptions even defy common sense. It is important to remember this when assessing applications of this methodology. Consider, for instance, Alwyn Young's work on the East Asian NICs,[v] where he claims that accumulation of capital and labor explains everything there is to about why these economies grew so rapidly. To assess this claim you have to find out whether the underlying assumptions, on which this conclusion is based, really hold. That means that you have to ask the following type of questions: Did perfect competition prevail? Were there no large firms with market power? No scale economics? Was technology freely available to everyone free of charge? Without answers to this these deeper questions growth accounting exercises cannot be used to draw conclusions about what drives growth. As pointed out by Robert Lucas,[vi] commenting on Young's findings, just observing the fact that input and output growth tends to go hand in hand explains nothing. Arguably, any theory of growth would be consistent with that!

With hindsight, the empiricists' attempt to base their work on the neoclassical (Solowian) theory of economic growth was less successful than initially conceived. What the empiricists failed to recognize, perhaps, was that the theory was not all geared towards the real world, but towards a totally artificial world ("perfect competition"), in which many of the growth inducing factors in the real economy had been eliminated by assumption. For the theoreticians this did the trick, since in this way were able to demonstrate that full employment was consistent with a market-oriented system of regulation. They did so at a high cost, however, because in contrast to the already existing post-Keynesian models of economic growth, the neoclassical model was unable to explain how the economy generates long run growth in GDP per capita. In fact, the only type of growth that the model could explain was transitional in character (on the path towards long run equilibrium), and associated with substitution of capital for labor. In the long run this type of growth was bound to cease. Any remaining growth of productivity would have to be classified as exogenous to the economic system. Economists often used the notion "manna from heaven" for such growth, which shows most clearly how they looked at it.

Alternative perspectives

I have used much space to explain how modern (neo-classical) growth theory (and its empirical applications) developed from the 1950s onwards. The justification for this is that it was so influential and dominated many peoples' thinking about growth for a long time. Indeed this perspective is still with us, although more controversial than it used to be. This is not to say that the post-Keynesian and neoclassical views were the only ideas about growth around at the time. For instance, alternatives to the neo-classical interpretations of events had for a long time been advocated by a diverse group of historically oriented economists and economically oriented historians. In lack of a better word I will classify these as "evolutionaries".[vii] I have chosen this wording to signal the importance that these people attached to the study of economic evolution in historical time and the drivers behind this process, particularly technological change which to a large extent was considered as endogenous.[viii]

The central contributor here was no doubt the Austrian-American economist Joseph Schumpeter. In contrast to the traditional emphasis in economics on capital accumulation, Schumpeter focused on innovation activities in firms as the driving force behind economic growth.[ix] Essentially, he sees innovations as "new combinations" of existing pieces of knowledge, whether drawn from science, engineering, market research, organizational experience or other sources, but with a view to commercial application. His concept of innovation is broad and goes beyond the mere invention of a new product or process. Successful innovating firms are assumed to benefit economically due to the temporary monopoly they get on the innovations they make. However, eventually the knowledge embodied in new innovations will diffuse to other firms and industries, and this will fuel growth further.

While Schumpeter particularly emphasized on deliberate innovation activities by firms, other writers in this tradition such as, for instance, Nicholas Kaldor[x], Bengt Åke Lundvall[xi] and Nathan Rosenberg[xii] have focused on the importance of learning as the source of technological progress and growth. Such learning occurs because people in their daily life, particularly at work, experience problems and – upon reflection – come up with new and improved solutions that increase productivity. It may originate in production, through investments and the subsequent application of new machinery, as the result of interaction with customers or suppliers or through organized links with other firms or organizations. Learning may also give rise to organized R&D of the type emphasized by Schumpeter, to some extent blurring the traditional distinction between innovation and diffusion of technology. Hence, in this literature, learning is increasingly analyzed as an interactive process, with feedback to and from organized R&D whether in the private or public sector. This has recently led researchers in this area to view a country's innovation and learning performance from a system perspective, focusing not only on the various institutions and organizations that take part in innovation and learning, but also on their mutual interaction. The concept "national system of innovation", used in several recent studies, reflects this perspective.[xiii]

The economic historian Alexander Gerschenkron pioneered the study of the international aspects of this process of innovation and learning. Some countries are at the technological frontier, while others lag behind. Although the technological gap between a frontier country and a laggard represents as "a great promise" for the latter (a potential for high growth through imitating frontier technologies); there are also various problems that may prevent backward countries from reaping the potential benefits to the full extent. Gerschenkron argued that if one country succeeds in embarking on an innovation-driven growth path others might find it increasingly difficult to catch up. His favorite example was Germany's attempt to catch up with Britain on century ago. When Britain industrialized, technology was relatively labor intensive and small scale. But in the course of time technology became much more capital and scale intensive, so when Germany entered the scene, the conditions for entry had changed considerably. Because this, Gerschenkron argued, Germany had to develop new institutional instruments for overcoming these obstacles, above all in the financial sector, "instruments for which there was little or no counterpart in an established industrial country".[xiv] He held these experiences to be valid also for other technologically lagging countries.

Moses Abramovitz,[xv] arguing along similar lines, has used the concepts "technological congruence" and "social capability" to characterize the situation for latecomers. The first concept refers to the degree to which leader and follower country characteristics are congruent in areas such as market size, factor supply etc. The second points to the various efforts and capabilities that backward countries have to develop in order to catch up, such as improving education, infrastructure and, more generally, technological capabilities (R&D facilities etc.). He explains the successful catch up of Western Europe vis a vis the USA in the post-war period as the result of both increasing technological congruence and improved social capabilities. As an example of the former he mentions how European economic integration led to the creation of larger and more homogenous markets in Europe, facilitating the transfer of scale-intensive technologies initially developed for US conditions. Regarding the latter, he points among other things to such factors as the general increases in educational levels and the rise in the share of resources devoted to public and private sector R&D. In a similar vein the failure of many so-called developing countries to exploit the same opportunities is commonly explained with reference to lack of technological congruence and missing social capabilities (f.i. education).

In a sense these alternative contributions paint a much bleaker picture of the prospects for catch-up than the traditional neoclassical theory discussed earlier. Catch-up is not something that can be expected to occur only by market forces left alone, but requires a lot of effort and institution-building on the part of the backward country. One main reason for this is no doubt that technology is generally viewed rather differently than in the standard neoclassical approach. Rather than something that exists in the public domain and can be exploited by anybody everywhere free of charge (the public good assumption), technological competence, whether created through learning or organized R&D, is commonly seen as deeply rooted in the specific capabilities of private firms and their networks/environments (including in many cases parts of the public sector).

The convergence controversy

For a long time empirical work on economic growth was dominated by measurement (attempts to measure productivity in different countries, sectors and industries) and growth-accounting exercises. However, as the empirical research agenda shifted from description towards explanation (of differences in growth between countries), researchers in this area started to supplement this descriptive work with econometric techniques (multivariate regression) to cross-country data sets with the purpose of distinguishing between the potential for catch-up and the various factors that determine to what extent this potential is actually exploited.[xvi] The potential for catch up was normally measured by the distance in productivity (or GDP per capita) between the country in question and the economically leading country of the sample (normally the United States). Other variables that were taken into account included differences across countries in rates of investment, educational attainment, R&D/innovation performance, openness to trade, size of government etc.

Initially most studies were confined to the developed market economies for which data were most easily forthcoming (the OECD countries). It was shown that among these countries, a process of catching up had taken place from the 1950s onwards, i.e., that the initially poorest countries in the area had grown much faster per capita than the economically leading country (the United States). Moreover, the differences in GDP per capita had steadily been reduced between the OECD countries, suggesting a tendency towards convergence towards a common level of GDP per capita for the area as a whole. It was also shown that these tendencies towards catch-up and convergence were much stronger when other conditioning variables were taken into account, indicating that the potential for catch-up/convergence was larger than what was actually realized. Many variables were found to contribute to this process, including – notably – investment, education and R&D/innovation performance. [xvii]

Although several of these studies were inspired by evolutionary views on technology and catching-up, the results could also be interpreted as supportive of the basic orthodox belief in the efficiency of markets, since convergence to a common level of productivity was indeed found to take place. However, this evidence turned out to be more controversial than initially conceived. For instance, it was pointed out by Bradford DeLong[xviii] in a paper from 1988 that the sample of OECD countries might be biased, since it consisted mainly of the countries in the global economy that had done reasonably well after the end of the second world war. He also presented some preliminary evidence suggesting that a similar tendency towards convergence could not be established for a more balanced sample. This pointed to the need for larger samples than just the countries of the OECD area, something that was made possible by the construction of new and larger data sets that were made available towards the end of the 1980s.[xix] Inspired by the work by Abramovitz and others on technology gaps and growth William Baumol and others[xx] applied regression models of the type just discussed to cross-country samples including up to 100 countries or more. The conclusion of this work was that while a tendency towards convergence could perhaps be established for the OECD countries in the post war period, and may be extending to some other countries as well, it does not hold for the world as a whole. In fact, many poor countries fail more or less completely to exploit the potential for catch-up, something that obviously does not go on well with the predictions of the traditional neoclassical theory. But the new evidence confirmed the previous finding of a considerable potential for catch-up by poorer countries, which, however, is not fully exploited due to lack of "social capability" and other factors (so-called "conditional convergence").

Hence, the evidence pointed towards a quite complex picture, with groups of countries with certain common characteristics doing rather differently. One interpretation of this evidence was that rather than global convergence what could be observed was a multitude of different "convergence clubs".[xxi] How should this be explained? Would it be possible to explain this diversity within a common theoretical framework, though necessarily more complex than the one that had dominated up to this point (that of Robert Solow)? This was the challenge confronting growth theorists towards the end of the 1980s.

The new orthodox theory

The problems that traditional neoclassical growth theory and empirical applications faced in explaining the observed patterns of long run growth in the world economy, and the emergence of other, competing approaches, led eventually to a search among neoclassical theorists for new models of growth. What they wanted was a formal model that continued to be based on the orthodox vision of the economy as a set of rational, maximizing agents (endowed with perfect information and foresight) interacting with one another but which yielded predictions that were consistent with what could actually be observed.

The central contributor here has been Paul Romer. There were in particular two aspects of the old neoclassical theory of economic growth that he wanted to improve on. First, he wanted a theory (or "model") that could explain long run growth without having to rely on the assumption of exogenous technological progress. Hence, he wished to endogenize technological progress. For this reason the theories that he helped to create are sometimes called "endogenous growth theories" (in contrast to the old neoclassical theory in which long run productivity growth had to be explained by exogenous factors). These theories are also sometimes dubbed "new growth" theories which is, of course, not very informative and – as times go by - not very accurate either. Second, he wanted this model to yield predictions that were consistent with the diversity of growth patterns that empirical research had found to exist, in particular the fact the poor countries did not catch up with the rich ones, but continued to stay poor (at least in relative terms).

In so doing he encountered the problem that we have already discussed at some length, namely that within the usual neoclassical framework, in the absence of exogenous technological progress, the assumptions of (a) positive, but decreasing returns to substitution of one factor (say, labor) by another one (say, capital) and (b) constant returns to a simultaneous increase in all factors, lead to a situation in which there is no incentive to further accumulation of capital per worker and, hence, no productivity growth. To avoid this outcome you have to allow for economies of scale in (i.e., increasing returns to scale) one way or another, so that accumulation can go on in spite of decreasing returns to factor substitution. This was of course not an entirely new idea, it had for instance been suggested by Nicholas Kaldor already in the 1950s. The reason why orthodox theorists avoided it for so long was probably that it was believed to be in conflict with their vision of a self-regulating market, since under such conditions, large firms would be more efficient than smaller ones, which, in the long run, might lead to some kind of monopoly. In such a situation, neoclassical theory would actually justify extensive intervention in the economy by the government, in contrast to what economists generally held to be the preferred economic policy stance.

However, Paul Romer showed that it was an easy way out of this, and that was to assume increasing returns to scale at the level of the country, and constant returns at the level of the firm, so that firms continue to operate as they were living within a world characterized by "perfect competition".[xxii] The idea was the simple one - suggested by among others Kenneth Arrow and Nicholas Kaldor two decades earlier - that the use of new forms of capital equipment leads to learning that may help improving new generations of machinery. The beneficiaries of this learning will be the users of these new generations of capital equipment. The firm(s) in which such learning occurs will also benefit, but not more than other users of this equipment. Hence, the learning process will leave the relative competitive position of firms unchanged, and will - therefore - not induce changes in their behavior vis a vis each other.

At the macro level, however, the consequences are different form those of the traditional Solow model. As pointed out, although learning cannot be appropriated by any single firm, all firms in a country are assumed to benefit collectively from it, in the forms of new and more productive machinery. The ever-increasing productivity of new generations of capital, caused by learning, checks the tendency towards decreasing returns to capital accumulation that would otherwise have led productivity growth so slow down and eventually - in the absence of exogenous technological progress - to cease altogether. Hence, because of learning, the capitalists will continue to find it profitable to invest in new machines. As a consequence growth will be sustained into the long run. This also implies that the forces that in the traditional Solow model were assumed to lead to convergence between rich and poor countries are no longer operative, because there is not any longer an inherent tendency for capital accumulation to slow down as the amount of capital per worker increases. Hence, rich countries may grow as fast as the poor ones, consistent with the apparent lack of convergence in the global economy.

To some extent this model did the trick, but it had a major disadvantage, it did not allow for technological progress to be caused by organized R&D within firms. Obviously, in a world with no unique benefits accruing to firms that invest in R&D, such as in the model discussed above, there will be no R&D either. To allow for technological progress to be caused by R&D, firms investing in R&D must – at least on average – receive an adequate return on these investments (as empirical research indeed suggests[xxiii]). But what is the economic mechanism generating such returns? The lion's share of the R&D investments are typically made early in the life of a product, often before it enters the market, and – in the case of success – paid back over the products life-time by keeping prices well above production cost. This, however, implies that the innovating firms have sufficient market power to keep prices at that level, in contrast to what is assumed to be the case in "perfect competition". Joseph Schumpeter, as we may recall, had explained this as the result of the temporary monopoly that innovating firms get on the innovations they make, which might be related to legal forms of protection (patents etc.), but also – and perhaps more commonly - to the fact that imitation in many cases is difficult, time-consuming and costly (because the total pool of knowledge and other assets necessary to produce and commercialize an innovation may not be very transparent or easily available on the market).

Basing himself partly on these ideas, Paul Romer suggested in later work an alternative theory, partly based on Schumpeterian ideas, in which both economies of scale and imperfect competition are assumed.[xxiv] In contrast to the previous model, in which technological progress was considered as a pure externality, like “manna from heaven”, this new approach models innovation as the outcome of deliberate efforts by firms which have sufficient market power to prevent the immediate and cost-less diffusion of their innovations to other firms and countries. However, Romper points out, every investment in R&D has to aspects, one specific related to a new product and process, that may be protected by patents, trade marks, secrecy or other means, and one general that contributes to the advance of scientific and technological knowledge in society as a whole, and which improves the capability to produce new innovations in the future. In this sense every new innovations stand on the shoulders of all previous ones. It is this continuous improvement in our capability to innovate that in this perspective prevents the decreasing returns to investments in innovative activity that would otherwise have set in.

Thus, in this second approach, long run economic growth is explained through the interplay of imperfect competition, which is necessary for R&D to take place, and spillovers (i.e., externalities) from these R&D investments on the general level knowledge in society and, hence, on the our capability to produce innovations in the future. The main difference between this framework and the previous one is that in this case it is the resources devoted to R&D - and the factors that influence the allocation of resources to this purpose - that determine economic growth, not capital accumulation in the traditional sense. Hence, from this perspective, the failure of many poor countries to catch up with the richer ones would among other things have to be explained by lack of resources devoted R&D (rather than physical capital).

These new theories do have interesting implication for policy. In the old framework, where productivity growth in the long run depended only on exogenous technological progress, policy by definition could not have a long-run impact. In these new models this is not longer so. Policies that impact on the propensities to invest in physical capital (the first type of model) or R&D/innovation (the second one) may raise growth permanently. Hence, as pointed put by Robert Lucas, from this perspective it is quite easy to conceive situations in which interventions by the government in the economy might have a significant, positive effects for the long-run performance of an economy.[xxv] However, it is difficult within such a framework to draw very firm conclusions on the use of specific policy instruments, since the appropriateness of these will depend on the characteristics of the country in question (country size, industrial structure, skill-composition of labor force etc.) and its relations with the global economy.[xxvi] For instance, one of the more controversial predictions from these theories is that large countries tend to benefit more from spillovers than others, and therefore is likely to grow faster. Moreover, if there is one economic activity characterized by extensive spillovers between firms (say, R&D-intensive industry) and another that is not (say, traditional industry) the theory suggests that large countries tend to specialize in the former and small countries in the latter. If true, and not only an artifact caused by, for instance, simplifications used in formal modeling, this implies that countries of different sizes may find themselves in quite different situations and, hence, would be inclined to chose quite different policy responses.

The evidence

These theoretical advances led to a surge of empirical work. As the new theory differs from the old one in important respects one might perhaps have expected that a new type of empirical work would have developed, focusing on new issues, using new data and applying new methods. This, however, has generally not been the case, or at least not until very recently.

What most applied researchers in this area have done is to follow the tradition from Cornwall, Baumol and others, i.e., applying regression models to cross-country data sets.

Ross Levine and David Renelt have summarized much of this work. What they did was to test the various factors that have been emphasized in the empirical literature in a systematic way in order to establish how robust the findings are for inclusion of other possible explanatory variables.[xxvii] This, it may be noted, is not a test of causality, important relationships may well be found to be fragile following this methodology. The principal finding of Levine and Renelt was that the most robust relationship is between growth and investment. Some support was also found for variables reflecting the scope for catch-up (proxied by GDP per capita gaps) and educational efforts. All other explanatory variables were found to be fragile, including a large number of policy variables, openness (defined in different ways) and political factors (such as democracy, stability etc.). [xxviii]

What is there to learn from this new generation of empirical research? Not very much I will argue. That investment is correlated with growth should come as no surprise. Indeed, this is something that would be consistent with most theories in this area, including those that consider investment as endogenous to the growth process, as some available evidence on time series data seems to suggest.[xxix] It is also worth noticing that the studies by Levine and others fail to include R&D and innovation, and thus throw little light on the mechanisms highlighted by the most recent generation of growth theories. However, the results from the empirical literature are useful in the sense that they urge us to use some caution when assessing the impact of policy on growth.

This lesson is especially relevant for those who believe that a so-called "correct" set of macro policies in combination with trade liberalization and deregulation is enough to foster development and growth, as argued, for instance, by the World Bank.[xxx] In fact, as pointed out above, there is very little scholarly support for such an interpretation of events. On the contrary it is clear that the governments in the most successful "catching-up" countries in the post-war period have all intervened extensively in the markets through various types of pro-active policies.[xxxi] Although the chosen policy set may have differed from one country to another, they have by and large performed the same function; to increase the share of national resources devoted to growth, and to steer these resources to the technologically most progressive parts of the economy. This is a recipe for high growth that would be consistent with several versions in the most recent generation of growth theory. The World Bank has nevertheless tried to argue that spectacular growth of these countries in the post war period these policies is not related to these policies, but as shown by several studies, these arguments do not sustain detailed scrutiny.[xxxii]

Another relevant strand of research attempts to measure private and social returns to R&D and innovation. This type work has gone on for a long time, independently of the developments in growth theory, but attracts a growing interest due to the recent changes in formal theorizing. Generally, these empirical exercises[xxxiii] tend to find high private returns to investments in R&D, about twice as high as for other types of investment. This, of course, runs counter to traditional neoclassical perspectives on investment, according to which returns to different types of investments should be equalized. Hence, one of the central issues in this area, which we will not venture into here, has been how these high private returns can be explained. However, high as these private returns may be, social returns are commonly found to be even higher, indicating important positive spillovers from R&D, especially when conducted in private firms. These findings suggest that from a social point of view, substantial underinvestment in (private sector) R&D is taking place, and that is an area where governmental intervention might be justified. These are, of course, results that concur very well with much recent theorizing in this area.

Recently, there have been some attempts to address these issues from a perspective that draws more explicitly on the advances in the growth literature. Central questions in this more recent literature are (a) to what extent the spread of new technology from the innovator to other firms, i.e., technology diffusion, is influenced by geographical, institutional and cultural boundaries, (b) whether country size matters for the degree of success in innovation (as suggested by some recent theorizing) and (c) what the most efficient carriers of technology diffusion are. Arguably, new technology may diffuse in many different ways; embodied in goods or services that make use new technology, through foreign direct investments by multinational firms or by imitative activities by domestic firms, drawing on a multitude of sources, as well as (necessary) complementary assets/capabilities. There is some theoretical backing for all of these but there has until recently been little if any evidence on their relative importance.

Although research in this area is still in early stage, the available evidence seems to indicate that diffusion of technology (knowledge spillovers) is hampered by distance and is generally easier and quicker within than across country borders.[xxxiv] There is also some evidence suggesting that returns to R&D investment are higher in large countries, consistent with some of the suggestions from recent advances in growth theory.[xxxv] On the last question, which carrier of technology is the most efficient, there is not much evidence yet. Some recent exercises point to R&D embodied in imports of goods and services as a very efficient way of transmitting new technology.[xxxvi] However, others, using essentially the same type of indicator, fail to reproduce these results, so the conclusion has to be that on this issue there is no firm evidence yet.[xxxvii] Finally, several recent studies indicate that there exist persistent differences across countries in their capacity to absorb foreign technology (education, infrastructure, technological capabilities etc.), and that this is a very important – probably the most important - factor explaining differences in growth and welfare across countries.[xxxviii]

Thus, the picture that emanates from this research is on which innovation and diffusion of technology plays an essential part for growth, and where diffusion, contrary to the assumption in the public good approach, is a difficult and costly process that requires a lot of effort not only by the firms themselves, but also by government at various levels. This is so because available evidence indicates that from a social point of view, there is considerable underinvestment in innovation of technology, and because research has shown that the diffusion of these innovations, which is essential for growth, depends on a number of capabilities that firms are unlikely to create to a sufficient extent if left to themselves. As for instance Joseph Stiglitz has pointed out, such coordination failure is a characteristic feature of industries operating under increasing returns. Arguably, the, at least until recently, seemingly very successful pro-active policies in catching-up countries concur very well with this perspective.

Conclusion

What the various perspectives on growth discussed in this essay have in common is that they contain two basic elements, a view on what drives growth and what form of regulation mechanism that is necessary to allow growth to go on.

On what drives growth, the dominant view has over the years been that growth is driven by accumulation of capital or "mechanization". This was the view of the classical political economists, including Karl Marx, post-Keynesian growth theorists such as Sir Roy Harrod or Evsey Domar (or Nicholas Kaldor for that sake), and – at least until very recently – all neoclassical theorists, including Robert Solow and Kenneth Arrow. From a historical point of view it is not difficult to understand why this view emerged as the dominant one. Clearly, during the so-called "industrial revolution" (whether it is appropriate to use the word revolution or if it should rather be seen as an evolutionary process) and the period that followed, mechanization was a vital ingredient.

It is perhaps less evident why this view should dominate economists' perceptions of the word for so long. In my view, it was the great achievement of Joseph Schumpeter (although he was not without forerunners) to break with this one-sided view and bring to the economists' attention a totally different argument about what drives growth, focusing on qualitative (i.e., innovation) rater than quantitative change. With hindsight it was probably no coincidence that Schumpeter's own professional career run parallel with the growth of science based industry, organized R&D and the development of various types of institutions and organizations relating to this process. It testifies to his qualities as an analyst of contemporary developments that he was able to grasp the full impact of these tendencies at such an early stage. In fact, it took a long time for science based industry and organized R&D to acquire the prominence it has today. Although pioneered in Germany at the turn of the century, it was only during the Second World War and the cold war that followed that these developments took off, and then primarily in the USA.

While initially most writers on growth shared a common conception of what were the drivers, this was not case for regulation. Since the beginning, discussions of regulation have been dominated by adherents of two diametrically opposite positions, that of "laissez faire" capitalism on the hand and state planning (or very extensive intervention) on the other, and the upper hand in the struggle has tended to change over time. Among the classical political economists, who all saw the economy from the mechanization perspective, some, such as Adam Smith and David Ricardo, were liberal free marketers, while Karl Marx believed that laissez faire was doomed to collapse, and had to be succeed by state planning. In the first half of this century, the view that a capitalist order ruled by free markets was basically unstable gained prominence as the crisis of the interwar period deepened, and led to the formulation of (the so-called post-Keynesian) growth theories advocating extensive state intervention in the economy. However, under the impact of the post-war boom and the cold war that followed, the liberal view again got the upper hand, as reflected in the neoclassical model of economic growth developed by Robert Solow and others.

The problem for this theory was not so much that it advocated free markets, which has been in fashion for a long time, but that it was based on an outdated understanding of what drives growth and how the capitalist system, including its institutions, works. It simply did not have any useful advice to give to people (such as policy-makers) who were concerned about the working of the economy. This is the reason why the works of Joseph Schumpeter and other economists with insights in how innovation and learning is shaped by – and shapes – the economy started to gain prominence again from the 1960s onwards, mainly through the writings of applied economists such as Christopher Freeman.[xxxix] Although Schumpeter himself was a devoted liberal, much of the work that based itself on his ideas came to focus on limitations to the working of markets, particularly with respect to innovation and diffusion of technology, and what the government - in countries on different levels of development - may do to improve on the economy in this respect.

What has happened in the area of formal theorizing in the last decade is that this agenda has been taken over by formal theorists. This has led to the creation of more complex models, incorporating technology and innovation, that explain growth in a better way than before. These models are also more open in the sense that many different outcomes are possible, depending on what the key assumptions are, and with a much greater room for policy. Many of these assumptions cannot be established on a priori grounds, at least not at the current state of formal theorizing in this area, but need to be verified through empirical research. This has led to a new agenda for empirical research that is both more meaningful and more interesting than what we had before. The most important contribution, I will argue, that empirical work can make to theoretical work is not to test assumedly "true" formal relationships but to raise the quality of the assumptions that theoreticians make use of.

While formal modeling in this area has greatly improved in terms of how technology is handled, other basic neoclassical features have retained, such as for example the idea of "the representative agent", i.e., that all agents in the economy are identical, rational, endowed with perfect information etc. This certainly runs counter to one of the most basic arguments of evolutionary reasoning; that agents are different (heterogeneous), that this creates diversity, which is what drives innovation. Hence, following this view, to explain diversity of growth patterns, one has to allow for heterogeneous agents, whether at the level of the individual entrepreneur, firm or nation state. This is an area where more work is needed, and which potentially could be of great importance for formal theorizing. However, to be able to respond to this challenge, empirical researchers have to go beyond empirical approaches that essentially consist of filtering out heterogeneity. Arguably, to get a firm grasp on heterogeneity, one will need more case-oriented research of the type undertaken in many other disciplines, as well as by the grand economic masters of the past, such as Adam Smith, Karl Marx, Alfred Marshall and Joseph Schumpeter.[xl]

Note:

[i] This chapter draws partly on an earlier paper, "Technology, Growth Evidence and Policy: Theory, Evidence and Interpretation", presented at the conference "Innovations, Policy and Society" in Oslo, December 4-5,1997, to appear in the Nordic Journal of Political Economy (NOPEC).

[ii] For an overview see Luigi Pasinetti, Growth and Income Distribution, Cambridge: Cambridge University Press, 1974

[iii] See Moses Abramovitz, Resources and Output Trends in the United States since 1870, American Economic Review 46, pp. 5-23 and Edward F. Denison, Why Growth Rates Differ: Post-War Experience in Nine Western Countries, Washington, D.C: Brookings Institution, 1967. A good overview of the research in this the area is to be found in Angus, Maddison, Growth and Slowdown in Advanced Capitalist Economies: Techniques of Quantitative Assessment, Journal of Economic Literature 25, pp. 649-698

[iv] Richard R. Nelson, Aggregate Production Functions and Medium-Range Growth Projections, American Economic Review 54:, pp. 575-606

[v] Alwyn Young, The Tyranny of Numbers: Confronting the Statistical Realities of the East Asian Growth Experience, Quarterly Journal of Economics 110, pp. 641-680

[vi] Robert E. Lucas, Making a Miracle, Econometrica, 61, pp. 251-272

[vii] Hence, I use the notion "evolutionary" in a broad sense. In a more narrow sense the notion is also used to characterize a specific class of formal economic models, which are inspired by the use of evolutionary models in biology. As in biology, these economic models operate with a set of competing "species", which in economics might be firms or technologies or both, and a selection environment which set the rules for their survival/growth. The central reference here is Richard R. Nelson and Sidney G. Winter, An Evolutionary Theory of Economic Change, Cambridge (Mass.): Harvard University Press, 1982. For an overview and

discussion of work in this area see Esben Sloth Andersen, Evolutionary Economics-Post-Schumpeterian Contributions, London: Pinter, 1994

[viii] It is worth mentioning that the emphasis on historical analysis and technological change was quite widespread among German economists around the turn of the previous century, particularly among the adherents of the so-called "historical school", and that Schumpeter - although he never regarded himself as belonging to any school – was probably influenced by this. For his own view on the "historical school" in Germany and the controversies that followed in its wake, see Joseph Schumpeter, History of Economic Analysis, Oxford University Press, New York, 1954, ch. IV.4

[ix] Joseph Schumpeter, The Theory of Economic Development, Cambridge, Mass.: Harvard University Press, 1934 (first German edition 1911)

[x] Nicholas Kaldor, A Model of Economic Growth, Economic Journal 67, pp. 591-624, and Nicholas Kaldor, Strategic Factors in Economic Development, State School of Industrial and Labor Relations, Cornell University, Itacha, New York, 1967. Similar ideas, though in a more neoclassical framework, were proposed by Kenneth Arrow, The economic implications of learning by doing, Review of Economic Studies 29, pp. 155-173

[xi] Bengt Åke Lundvall, "Innovation as an interactive process: from user-producer interaction to the national system of innovation", in Giovanni Dosi et al., Technical Change and Economic Theory, Pinter, London, 1988, pp. 349-369

[xii] Nathan Rosenberg, "Learning by Using" in Nathan Rosenberg, Inside the black box: Technology and Economics, Cambridge University Press, Cambridge, 1982, pp. 120-140

[xiii] See Bengt Åke Lundvall, ed, National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning, London: Pinter, 1992, Richard R. Nelson (ed) National Innovation Systems: A Comparative Analysis, Oxford: Oxford University Press, 1993, Charles Edquist (ed), Systems of Innovation: Technologies, Institutions and Organizations,

London: Pinter, 1997 and Christopher Freeman, The national System of Innovation in Historical Perspective, Cambridge Journal of Economics 19, pp. 5-24

[xiv] Alexander Gerschenkron, Economic Backwardness in Historical Perspective, Cambridge(Mass.): The Belknap Press, 1962, p. 7

[xv] Moses Abramovitz, "The Origins of the Post-war Catch-Up and Convergence Boom", in Jan Fagerberg, Bart Verspagen. and Nick von Tunzelman (eds.), The Dynamics of Technology, Trade and Growth, Edward Elgar, Aldershot, 1994, pp. 21-52

[xvi] The pioneer in much of this was John Cornwall who regressed variables assumed to reflect the scope for catch-up, investment and endogenous technological progress (the so-called Verdoorn law) on GDP growth for a sample of OECD countries, see John Cornwall, Diffusion, Convergence and Kaldor's Law, Economic Journal 86, pp. 307-314. Since then a great number of studies of this type, including the scope for catch-up and some other conditioning variables as determinants of GDP/productivity growth has been published, for an overview see Jan Fagerberg, Technology and International Differences in Growth Rates, Journal of Economic Literature 32, pp. 1147-1175. In the most recent literature, such regression models have for some reason been dubbed "Barro-regressions" after Robert Barro, an economist who did not pioneer this type of research, but picked it up at a rather late stage.

[xvii] Although relevant, many applied studies did not include R&D or innovation performance, often with reference to lack of data (although this was not always entirely justified). For an exception to this trend see Jan Fagerberg, A Technology Gap Approach to Why Growth Rates Differ, Research Policy 16, pp. 87-99

[xviii] Bradford De Long, Productivity Growth, Convergence and Welfare, American Economic Review 78, pp. 1138-1154

[xix] Robert Summers and Alan Heston, The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988, Quarterly Journal of Economics 106, pp. 327-368

[xx] William J. Baumol, Sue Anne Batey Blackman and Edward N. Wolff, Productivity and American Leadership: The Long View, MIT Press, Cambridge, Mass.,1989

[xxi] Ibid.

[xxii] See Paul M. Romer, Increasing Returns and Long-Run Growth, Journal of Political Economy 94, pp. 1002-1037

[xxiii] See the survey by Zvi Griliches, The Search for R&D Spillovers, Scandinavian Journal of Economics 94, pp. S29-S47

[xxiv] Paul M. Romer, Endogenous Technological Change, Journal of Political Economy 98. Pp. S71-S102. Models along similar lines were also suggested by a number of other authors, see for instance Pierre Aghion and Peter Howitt, A Model of Growth Through Creative Destruction, Econometrica 60, pp .323-351 and Gene M. Grossman and Elhanan Helpman (1991) Innovation and Growth in the Global Economy, The MIT Press, Cambridge(Mass.), 1991.

[xxv] Robert E. Lucas, On the Mechanisms of Economic Development, Journal of Monetary Economics 22, pp. 3-42

[xxvi] See Gene M. Grossman and Elhanan Helpman, op.cit.

[xxvii] The method consists of selecting a set of basic variables, which always are included in the regression. Other possible variables are included one by one and the sensitivity of the result is then tested by including up to three other variables drawn from a large pool of possible explanatory factors. If the variable is always significant, the relationship is termed "robust". If it is insignificant in at least one case it is considered as "fragile". See Ross Levine and David Renelt, A Sensitivity Analysis of Cross-Country Growth Regressions, American Economic Review 82, pp. 942-963

[xxviii] In a later study by Robert King and Ross Levine, Finance and Growth: Schumpeter Might Be Right? Quarterly Journal of Economics 108, pp. 717-737, the level of financial

development of the country was added to the list of robust variables.

[xxix] Christopher D. Carrol and David N. Weil, Saving and Growth: A Reinterpretation, NBER Working Paper, No. 4470, National Bureau of Economic Research, Cambridge(Mass.), 1993

[xxx] World Bank, The East Asian Miracle. Economic Growth and Public Policy, Oxford University Press, New York, 1993

[xxxi] See, for instance, Robert Wade, Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization, Princeton: Princeton University Press, 1990

[xxxii] For instance, in a World Bank report on the East Asian miracle from 1993 (World Bank, op.cit.) the bank argues that 60-90 % of productivity growth of the so-called high-performing Asian economies can be explained by accumulation and thus that other "unconventional" factors were of relatively little importance. I have shown elsewhere (Aadne Cappelen and Jan Fagerberg, East Asian Growth: A Critical Assessment, Forum for Development Studies, No. 2, 1995, pp. 175-195) that this conclusion is not warranted. In fact, the models applied by the World Bank predict very poorly for the fast-growing countries of Asia! There are also other attempts in the report to prove the case; Dani Rodrik shows these not more convincing. See Dani Rodrik, "King Kong Meets Godzilla: The World Bank and the East Asian Miracle," in Albert Fishlow et al., ed., Miracle or Design? Lessons from the East Asian Experience, Policy Essay No. 11, Washington, DC.: Overseas Development Council, 1994, pp. 13-53

[xxxiii] See the survey by Zvi Griliches, op. cit..

[xxxiv] See Adam B. Jaffe, Real Effects of Academic Research, American Economic Review 79, pp. 957-970, Adam B. Jaffe, Manuel Trajtenberg and Rebecca Henderson, Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, Quarterly Journal of Economics 108, pp. 557-598 and Per B. Maurseth and Bart Verspagen, "Europe: on or several

systems of innovation" in Jan Fagerberg, Paolo Guerrieri and Bart Verspagen (eds.), The economic challenge for Europe: Adapting to innovation-based growth, Aldershot: Elgar, 1999, pp.

[xxxv] See David T. Coe and Elhanan Helpman, International R&D Spillovers, European Economic Review 39, pp. 859-887

[xxxvi] The conclusion, then, would be that foreign R&D embodied in imports is the primary source of growth in most countries, particularly the developing ones, and that openness to trade is what is required if a country is going to benefit from the global process of innovation and diffusion. See David T. Coe and Elhanan Helpman, op.cit., and David T. Coe, Elhanan Helpman and Alexander Hoffmaister , North-South R&D Spillovers, Economic Journal 107, pp. 134-149

[xxxvii] Maury Gittleman and Edward N. Wolff, R&D Activity and Cross-Country Growth Comparisons, Cambridge Journal of Economics 19, pp. 189-207

[xxxviii] See Maury Gittleman and Edward N. Wolff, op. cit., and Jonathon Eaton. and Samuel Kortum, Trade in Ideas - Patenting and Productivity in the OECD, Journal of International Economics 40, pp. 251-278. Bart Verspagen argues that failure to take into account such differences may in fact explain some of the conflicting evidence reported in the literature on the importance of other carriers of technology diffusion. See Bart Verspagen, Estimating International Technology Spillovers Using Technology Flow Matrices, Weltwirtschaftliches Archiv 133, pp. 226- 248_

[xxxix] See Christopher Freeman, The Economics of Industrial Innovation, Harmondsworth: Penguin, 1974

[xl] This does, of course, not invalidate the use of other methods that are currently in more use in economics. Arguably, empirical work will need to proceed at several levels, not in isolation, but in interaction.